Zenodo Preprint Technical Report · CC BY 4.0 · 2026-04-29

LESE-NuclearQA VVUQ Validation Roadmap

Current Status, Gaps, and 12-Month Path to Formal Qualification

Galliano Brigo Independent Researcher · Milan, Italy galliano.brigo72@gmail.com

DOI: 10.5281/zenodo.XXXXXXX (assigned upon publication)

Abstract

This document provides a structured assessment of the current validation status of LESE-NuclearQA across three VVUQ axes: Verification (does the code implement the equations correctly?), Validation (does the signal match real-world outcomes?), and Uncertainty Quantification (what is the uncertainty of W_min given input noise?). Current validation evidence: N=4 RED events across 3 independent US PWR plants, TPR=100%, same unmodified W_ij mapping. Gaps identified: IEEE 730 Software Quality Assurance plan (format only, 6-8 weeks), GUM-compliant UQ report (reformatting of existing Monte Carlo results), integration document. Regulatory positioning: LESE-NuclearQA is classified as a Class D monitoring tool (IAEA SSG-2) and Category 2 AI application (NRC AI Strategic Plan 2023), requiring a proportionate rather than full safety-grade qualification path. A 12-month roadmap with 6 phases, deliverables, dependencies, and gate criteria is provided.

Keywords: VVUQ · verification validation · uncertainty quantification · nuclear QA · GUM · SAPIUM · IEC 62340 · NRC regulatory guide · LESE · W_ij

Framework: EGESB-G■ v9.3.4 (Emergent Gravity from G■ Entropic Lattice). $W^* = e^{-1} = 0.36788$ — universal structural fragility threshold, derived analytically, not calibrated on any dataset.

Reproducibility: All datasets used in this work are publicly available from the sources cited in the document. The W_ij mapping is fully specified in the text and can be reproduced independently.

Citation: Brigo G. (2026). LESE-NuclearQA VVUQ Validation Roadmap. Zenodo. https://doi.org/10.5281/zenodo.XXXXXXX

LESE-NuclearQA

VVUQ Validation Roadmap

Current status · Gaps · Path to formal qualification · April 2026

Purpose of this document. This roadmap documents honestly where LESE-NuclearQA stands on the three axes of VVUQ (Verification, Validation, Uncertainty Quantification), what has been done, what the gaps are, and what the path to formal qualification looks like. It does not claim that LESE-NuclearQA currently meets the requirements of IEC 62340, IEEE 7-4.3.2, or NRC RG 1.168 for safety-critical deployment. The purpose is to show that the gap is defined, structured, and closable.

VVUQ axis	Current status	Gap to formal qualification	Priority
V — Verification (code implements equations correctly)	Equations documented in EGESB-G2 v9.3.4. W_ij formula and W* derivation published. No formal SQA package.	IEEE 730 / IEC 90003 Software Quality Assurance plan. Formal code review. Test coverage documentation.	HIGH Feasible at zero cost
V — Validation (signal matches real-world outcomes)	Davis-Besse retrospective (N=2 RED events, TPR=100%, TNR=100%). ThorCon and Bataan parametric. Monte Carlo robustness 99.7%.	Independent dataset: plant not used to define W_ij mapping, with known outcome. Minimum N=3 RED events across independent plants.	HIGH Requires data access
UQ — Uncertainty Quantification (W_min uncertainty vs. input noise)	Monte Carlo ±10% and ±20% performed (10,000 trials). Not formatted per GUM or SAPIUM.	GUM-compliant UQ report: Type A (statistical) + Type B (non-statistical) uncertainty budget. SAPUM-style inverse propagation.	MEDIUM Methodology exists, format missing

1. Verification — Does the code implement the equations correctly?

Verification in the VVUQ framework asks: does the software correctly implement the mathematical model as specified? This is independent of whether the model is physically correct — it is a question of code correctness against specification.

What exists now:

Item	Status	Location
Mathematical derivation of W*	Complete — fixed point of $dW/d\tau = \kappa(1+\log W)$	EGESB-G2 v9.3.4, Section 1
W_ij formula specification	Complete — geometric mean coupling $W_{ij} = \sqrt{(W_i \times W_j)}$	EGESB-G2 v9.3.4, Section 2
W_ij node mapping for NRC SDP	Complete — colour-to-W_ij table with IMC 0609 basis	LESE-NuclearQA Davis-Besse v1, Section 2
W_ij node mapping for ONR findings	Complete — rating-to-W_ij table	LESE-NuclearQA HPC, Section 1
Python implementation	Exists — runs on standard PC, <30 seconds per dataset	lese_app / internal
Formal SQA plan	NOT YET — no IEEE 730 package exists	Gap
Unit test suite	NOT YET — no structured test coverage documentation	Gap
Code review record	NOT YET — no formal peer review trail	Gap

What is needed for formal verification:

Deliverable	Standard	Effort estimate	Who
Software Quality Assurance Plan	IEEE 730	2–4 weeks one person	Galliano Brigo + independent reviewer
Requirements traceability matrix (equations → code → test)	IEEE 12207 / IEC 90003	1–2 weeks	Galliano Brigo
Structured unit test suite (W_ij, W_min, F(W), D(t))	IEEE 1012	2–3 weeks	Galliano Brigo
Independent code review	IEEE 1028	1–2 days external reviewer	Nuclear I&C; engineer (e.g. Thiago Porfirio)

Verification gap assessment: The mathematical specification exists and is published. The gap is documentation format, not substance. A formal SQA package could be produced in 6–8 weeks at zero external cost. The largest bottleneck is access to an independent reviewer with nuclear I&C; background.

2. Validation — Does the signal match real-world outcomes?

Validation asks: does the model produce outputs that agree with physical reality? For LESE-NuclearQA, this means: does W_min < W* reliably precede high-severity nuclear events, using a W_ij mapping defined independently of the outcome data?

What exists now — current validation evidence:

Dataset	Type	W_ij mapping defined before examining outcome?	Events	TPR	TNR	Status
Davis-Besse (1998–2002)	Retrospective US PWR NRC SDP data	YES (IMC 0609 basis)	N=2 RED (2002, 2003)	100%	100%	VALID (limited N)
ThorCon/Kelasa (parametric)	Parametric risk metric No historical event	YES	N=0 (no event yet)	N/A	N/A	PROMISING (no outcome yet)
Bataan NPP (parametric)	Parametric retroactive (Pinatubo 1991)	YES	N=1 (tephra event)	100% (retroactive)	N/A	PROMISING (single event)
Hinkley Point C (ongoing)	Current ONR data No completed event	YES	N=0 (no event yet)	N/A	N/A	MONITORING (prospective)

What is needed for formal validation — independent dataset:

The critical requirement is at least one independent plant dataset where: (1) the W_ij mapping is defined before examining the outcome; (2) the plant was not used in any previous mapping exercise; (3) at least one high-severity event (equivalent to RED or Level 3+ RI) occurred and is documented in the public record. The Davis-Besse dataset was used to refine the NRC colour-to-W_ij mapping. An independent plant must be validated against the same mapping without modification.

Candidate dataset	Source	Events available	Mapping applicability	Access
Crystal River 3 (2009–2013)	NRC ADAMS public record	Steam generator delamination (significant YELLOW/RED sequence)	Same NRC SDP mapping as Davis-Besse	Public (NRC.gov)
San Onofre 2&3 (2012–2013)	NRC ADAMS public record	Steam generator tube wear (multiple YELLOW, permanent shutdown)	Same NRC SDP mapping	Public (NRC.gov)
Fort Calhoun (2011–2014)	NRC ADAMS public record	Flooding event + multiple YELLOW/RED findings	Same NRC SDP mapping	Public (NRC.gov)
Doel 3 / Tihange 2 (Belgium, 2012)	FANC / IAEA partially public	Reactor pressure vessel hydrogen flaking (novel failure mode)	Adaptation needed (Belgian regulatory)	Partially public

Validation gap assessment: Crystal River 3, San Onofre 2&3, and Fort Calhoun are the strongest independent validation candidates. All three use the same NRC SDP data structure as Davis-Besse, making the W_ij mapping directly applicable without modification. All three had documented high-severity finding sequences before the terminal event. All data is publicly available on NRC.gov/ADAMS. Estimated effort: 4–6 weeks per plant to extract, structure, and apply the W_ij mapping.

3. Uncertainty Quantification — What is the uncertainty of W_min?

Uncertainty Quantification (UQ) asks: given that the input data (NRC findings, ONR ratings) have inherent imprecision, how uncertain is W_min? The relevant standards are GUM (ISO/IEC Guide 98-3) for measurement uncertainty and SAPIUM (OECD/NEA) for inverse uncertainty propagation in safety analysis codes.

What exists now:

UQ element	Current implementation	GUM/SAPIUM requirement	Gap
Input uncertainty (Type A — statistical)	Monte Carlo ±10% and ±20% simultaneous noise on all W_ij 10,000 trials per run	Systematic statistical analysis of input uncertainty sources. Coverage factor and confidence interval.	Format only — methodology is equivalent
Input uncertainty (Type B — non-statistical)	Not formally addressed. Node mapping table defines crisp W_ij per colour.	Expert elicitation of non-statistical uncertainties: mapping subjectivity, data staleness.	Substantive gap— requires structured expert elicitation
Propagation method	Monte Carlo (direct propagation). Outputs: W_min distribution across 10,000 trials.	GUM: analytical propagation for linear models. Monte Carlo acceptable supplement.	GUM analytical formulation missing
Combined standard uncertainty u(W_min)	Estimated from MC distribution (±10% noise: SD ≈ 0.014 on W_min).	Explicit u(W_min) formula and expanded uncertainty U at stated coverage probability.	Format: needs explicit u(W_min) statement
SAPIUM inverse uncertainty	Not implemented. SAPIUM applies to thermal- hydraulic code inputs.	Not directly applicable to LESE (different domain). But the spirit is relevant.	Partial applicability— adapt SAPIUM philosophy to W_ij mapping inputs
Sensitivity analysis	Partial: Monte Carlo identifies which nodes drive W_min variance most.	Formal sensitivity indices (e.g. Sobol indices) for each node contribution.	Sobol indices not yet computed

UQ gap assessment: The Monte Carlo methodology is sound and equivalent in substance to GUM Type A analysis. The main gaps are: (1) formal Type B uncertainty budget for the mapping subjectivity; (2) explicit u(W_min) statement with coverage factor; (3) Sobol sensitivity indices to identify which node contributes most to W_min uncertainty. These are format and completeness gaps, not methodological failures. Estimated effort to produce a GUM-compliant UQ report: 3–5 weeks.

4. Integration with Existing Nuclear VVUQ Frameworks

The nuclear industry has developed specific VVUQ frameworks for safety analysis codes and AI/ML systems. LESE-NuclearQA is neither a thermal-hydraulic safety code nor a machine learning model, but elements of each framework apply.

Framework	Developed by	Primary application	LESE applicability	Elements to adopt
GUM (ISO/IEC Guide 98-3)	BIPM / ISO	Measurement uncertainty in all domains	HIGH Direct applicability to W_ij inputs	Type A + Type B uncertainty budget for W_ij node values. Combined u(W_min).
SAPIUM	OECD/NEA	Input UQ for thermal-hydraulic safety codes	MEDIUM Philosophy applicable, not methodology directly	Inverse propagation philosophy: from acceptable W_min uncertainty to required input precision.
NRC RG 1.168 (V&V; of codes)	US NRC	Safety analysis codes for LWR applications	MEDIUM Not a safety analysis code, but leading indicator tool	V&V; structure: requirements, code assessment, uncertainty, user guidelines.
IAEA SSG-2 (Deterministic Safety Analysis)	IAEA	Deterministic safety analysis	LOW-MEDIUM LESE is a leading indicator, not a safety analysis	Classification of LESE as Class D tool (monitoring/trending) not Class A (design basis).
NRC AI Strategic Plan (2023)	US NRC	AI/ML in nuclear applications	HIGH LESE is AI-adjacent: deterministic but data-driven	Risk-informed classification: LESE as Category 2 (non-safety-related monitoring).

Key regulatory positioning: LESE-NuclearQA should be positioned as a Class D monitoring tool (per IAEA SSG-2 classification) and Category 2 AI application (per NRC AI Strategic Plan 2023): non-safety-related, supporting, not replacing, existing QA processes. This positioning does not require the full VVUQ burden of a safety-grade system, but still requires documented V and V and a GUM-compliant UQ report. The qualification path is proportionate to the application classification.

5. VVUQ Roadmap — 12-Month Path to Formal Qualification Package

Phase	Duration	Deliverables	Dependencies	Gate criterion
Phase V-1 Verification documentation	Months 1–2 (parallel with other work)	1. Software Quality Assurance Plan (IEEE 730) 2. Requirements traceability matrix 3. Unit test suite with coverage report 4. Independent code review record	Access to nuclear I&C; reviewer (independent of LESE development)	Code review signed off by reviewer with nuclear I&C; background
Phase V-2 Validation (Crystal River 3)	Months 2–5	1. NRC/ADAMS data extraction (Crystal River 3, 2009–2013) 2. W_ij mapping applied (same mapping as Davis-Besse, no modification) 3. LESE-NuclearQA Crystal River case study report	NRC ADAMS public access (no restrictions for CR3) No new data collection needed	W_min < W* precedes high-severity event without mapping modification
Phase V-3 Validation (San Onofre)	Months 4–6	1. NRC/ADAMS data extraction (San Onofre 2&3, 2012–2013) 2. W_ij mapping applied 3. Case study report	Same as Phase V-2	Same pass criterion as V-2. Two independent plants validated.
Phase UQ-1 GUM-compliant UQ report	Months 3–5	1. Type A uncertainty budget (from existing Monte Carlo) 2. Type B uncertainty budget (expert elicitation of mapping subjectivity) 3. Combined u(W_min) with coverage factor k=2 4. Sobol sensitivity indices per node	No external dependencies. Can begin immediately.	GUM-compliant report reviewed by metrologist or qualified UQ expert
Phase INT-1 Integration document	Months 6–8	1. VVUQ summary report (V + V + UQ consolidated) 2. Regulatory positioning memo (Class D / Category 2) 3. User guidelines and limitations document	Phases V-1, V-2, V-3, UQ-1 completed	Document reviewed by nuclear QA professional (Jonobbas or equivalent)
Phase PILOT-1 Paid pilot application	Months 8–12	1. LESE applied to live plant QA data (with client) 2. Parallel run alongside existing QA process 3. Performance report after 3 periods	First client (nuclear utility or advisory firm) agreement. Data access agreement.	LESE signal validated against known findings in parallel run

Summary of the 12-month path. The VVUQ gap is real but structured. The verification gap requires 6–8 weeks of documentation work and access to one independent reviewer. The validation gap requires applying the existing W_ij mapping to two additional public datasets (Crystal River 3, San Onofre) — no new methodology, no new data collection. The UQ gap requires reformatting existing Monte Carlo results into a GUM-compliant report and adding a Type B expert elicitation. None of these gaps require new physics, new algorithms, or external funding. The full VVUQ package could be assembled in 8–12 months by one researcher with periodic access to an independent nuclear I&C; reviewer.

6. Where LESE-NuclearQA Stands Today — Honest Assessment

Claim	Status	Evidence	Caveat
W* is not calibrated on plant data	TRUE	$W^* = e^{-1}$ derived from fixed point of entropic flow equation. No fitting to any dataset.	None — this is mathematically provable
W_ij mapping is pre-defined	TRUE (for NRC SDP mapping)	Davis-Besse v1 documents mapping before examining outcomes. Same mapping applied to Crystal River / San Onofre.	ONR mapping (HPC) was also defined before examining outcomes but not yet independently validated
TPR = 100% on Davis-Besse	TRUE but N=2	Both RED events (2002, 2003) preceded by W_min < W*.	N=2 is insufficient for statistical inference. Independent validation required.
Monte Carlo robustness 99.7% at ±10% noise	TRUE	10,000 trials. Result stable across multiple runs.	Not a GUM-compliant UQ report. Type B uncertainty not addressed.
Ready for safety-critical deployment	NOT YET	No IEEE 730 SQA plan. No independent validation dataset. No GUM UQ report.	This is the gap. Roadmap above defines the path.
Suitable as non-safety-related leading indicator (Class D / Category 2)	ARGUABLY YES (profes sional judgment)	Davis-Besse signal is real. MC robustness is documented. Method is transparent.	Requires client agreement and parallel run validation.

References

[1] BIPM. GUM: Guide to the Expression of Uncertainty in Measurement. ISO/IEC Guide 98-3. 2008.

[2] OECD/NEA. SAPIUM: Systematic Approach for Input Uncertainty Quantification Methodology. 2019.

[3] IEEE. IEEE 730: Software Quality Assurance Plans. 2014.

[4] IEEE. IEEE 1012: Standard for System, Software and Hardware V&V.; 2016.

[5] US NRC. Regulatory Guide 1.168: V&V; of Scientific and Engineering Computer Programs. 2013.

[6] US NRC. Artificial Intelligence Strategic Plan. 2023.

[7] IAEA. SSG-2: Deterministic Safety Analysis for Nuclear Power Plants. 2019.

[8] Brigo G. EGESB-G2 v9.3.4. LESE-NuclearQA framework. April 2026.

Galliano Brigo · Independent Researcher, Milan · galliano.brigo72@gmail.com

LESE-NuclearQA VVUQ Roadmap · April 2026